

UNITED STATES PATENT APPLICATION

of

Aleksandar P. Radulovic

for

**PRIVATE IP COMMUNICATION
NETWORK ARCHITECTURE**

BACKGROUND OF THE INVENTION

Related Applications

This application is a continuation-in-part of application serial no. 08/599,238, filed February 9, 1996, entitled VOICE INTERNET TRANSMISSION SYSTEM to Wilkes et al. and serial number 08/585,628, filed January 16, 1996, entitled FACSIMILE INTERNET TRANSMISSION SYSTEM to Wilkes et al., each of which are incorporated herein by reference.

The Field of the Invention

10 The present invention relates generally to multimedia communication networks.
11 More particularly, embodiments of the present invention relate to an improved Voice over
12 Internet Protocol (VoIP) network that provides for increased data stream throughput for
13 video/voice/data via a private Internet Protocol (IP) communications network with
14 associated communications components.

Description of Related Art

To understand the need for the improved design of the present invention requires a brief overview of the current solutions that need to be replaced, the available technology, and the areas of anticipated future growth. As none of the existing VoIP standards support all necessary telephony signals and messages, vendors develop special proprietary messages and controls to allow important features. Therefore, the current market products do not interact at the capacity required to offer carriers separate high performance, scalable, and reliable building blocks to build viable nationwide VoIP networks.

1 The public switched telephone network (PSTN) was built over many decades to
2 accommodate basic telephony communications. As a result, other types of signals, such as
3 data, video, and fax are presently formatted inefficiently to fit into the framework of the
4 voice structure in the PSTN network. Presently, the telecommunication industry is
5 undergoing a dramatic reorganization and reconnection. As newer technologies mature and
6 develop, their deployment base, cost, and reliability approach the necessary size, expense,
7 and dependability levels required to replace the old public switch technology. There is also
8 a significant amount of engineering effort directed towards the selection of elements and
9 components built for the PSTN that could be reused by a telephone network successor.
10 Despite these changes, the infrastructure of the standard publicly switched telephone
11 network (PSTN) has remained virtually the same. What is needed is a protocol and
12 architecture that allows simple interfacing between new and legacy components.

13 The legacy PSTN supplies reliable and simple voice communication via an analog
14 data transmission medium, but is not well suited for modern digital communication. While
15 the public infrastructure actually consists of many different networks and technologies,
16 much of the system can be characterized as having two inherent shortcomings. First, the
17 reserved bandwidth in most voice calls is idle for much of the call duration, yet it is
18 unavailable to carry other traffic, creating a systemic inefficiency. Second, the only
19 intelligence in the system is found in the carrier's routing and control logic resident in the
20 switches and control network. Basically analog local loops are connected to a local class
21 five switch, which connects to other backbone and local switches through two separate
22 networks. One network of inter-machine trunks carries the actual media traffic in the form
23 of 64kbs time division multiplexed (TDM) streams. A separate packet switched network
24 carries call signaling, and control instructions using the SS7 protocol. As such, much of the

1 logic for call connection and routing is resident in the switches, but additional logic for
2 services, such as 800 number services, is drawn from Service Control Point databases on the
3 SS7 network. When a call is placed on this network, instructions sent over the control layer
4 allocate physical resources (ports and bandwidth) in the transport layer, creating a private
5 channel of fixed bandwidth that is maintained for the duration of the call. In essence, this is
6 a system that is highly adapted to a limited, historic set of functions, and which is not readily
7 adaptable to new types of services or a wider range of media inputs. What is needed is a
8 distributed architecture that allows data to be transmitted via a variety of message types
9 optimized for the data being sent and for the network being used.

10 In contrast to the PSTN, the Internet supplies reliable and rapid computer data
11 transmission without the added burden of long-distance charges. Originally developed by
12 the government to facilitate communication in adverse conditions, the DARPA project
13 consisted of a computer network that did not rely on any single node or cable for its
14 existence. DARPA was specifically developed to provide multiple pathways for
15 communication to flow from a source to a destination. Data can thus be routed along a large
16 variety of pathways, successful transmission is not dependent on any one single pathway for
17 a majority of the message to be successfully delivered. The successor to the DARPA project
18 is the popular and widely used Internet. Transmission of analog or voice data via the
19 Internet is viable because voice data can be digitized and the Internet is a global
20 transmission medium, which substantially duplicates the area covered by the PSTN. An
21 even greater advantage is the fact that Internet access generally includes all data
22 transmission fees in the base cost unlike the PSTN were the base cost only includes
23 connectivity and the user pays additional fees for data transmission, such as long distance
24 calls. Presently available PSTN systems cannot supply high connectivity without adding

1 unreasonable restrictions. Such systems should also supply support for multi-media and
2 variable message types. What is needed is a protocol and architecture that takes advantage of
3 the Internet's high connectivity, natural command and control infrastructure, multimedia
4 support, and uses low cost and low complexity internet-scalable devices

5 The standards organizations do not keep pace with modern technological
6 developments, due in part to the fact that the standards organizations are very political and
7 the increasing speed at which products are developed in the "Internet economy". Generally
8 a selected protocol tends to give technological preference to one vendor over another, so the
9 various vendors participating in the standards committees are naturally at odds with
10 proposals presented by other vendors. This slows the progress, development, and the
11 performance of the standards eventually implemented. What is needed is a truly efficient,
12 interoperable and carrier grade protocol for use over a packetized network. The standards
13 organizations, in general, and the VoIP market, in particular, are fragmented with many
14 different protocols that compete and overlap. Presently, the two major VoIP protocols
15 competing in the carrier market space are H.323 and SIP.

16 Developed originally for the transfer of multi-media signals over non-reliable
17 networks (such as LAN), H.323 has been transformed in an attempt to meet the needs of a
18 true carrier grade network. Although it is an approved standard by the International
19 Telecommunication Union (ITU), the H.323 protocol faces significant opposition in the
20 market due to its high complexity and lack of complete carrier grade support. In theory,
21 H.323 should enable users to participate in the same conference even though they are using
22 different videoconferencing applications. But it's too early to say whether such adherence
23 will actually result in interoperability, even though most videoconferencing vendors have
24 announced that their products will conform to H.323. What is needed is a packetized

communication protocol that provides carrier grade support through simple compatible building blocks.

A much simpler and modern protocol, SIP places special emphasis on SS7 support. As the SIP protocol is still secondary in the industry to H.323 in terms of deployment and support by vendors, the future of the SIP protocol is unclear. What is needed is a communication protocol that supplies an open interface to existing and emerging standards, such as SS7 and H.323.

As attention to VoIP rises, there is an increasing collection of vendors in the market. Many of these vendors have developed products that work well in labs or small-scale installations; however, they are completely unsuitable for a carrier grade network requiring a call capacity of many millions of calls per hour. Obtaining carrier grade performance should be one of the driving forces in the design of any new network component or network architecture. What is needed is a network architecture that designs each component, which could cause a bottleneck, as a distributed application, thereby enabling replication of the resource to enhance overall network capacity.

WORKMAN, NYDEGGER & SEELEY'S OFFICE

1

2 **SUMMARY OF THE INVENTION**

3

4 The present invention has been developed in response to the current state of the art
5 and, in particular, in response to these and other problems and needs that have not been fully
6 or completely solved by currently available communication networks. Processing power
7 (i.e., computers along with the internet infrastructure and data networking), has reached a
8 level where the technology implemented in the IP world may be utilized as a catalyst for
9 change in the telecommunications industry. For example, an Internet router is capable of
10 routing many times more data than a traditional telecom switch that costs several orders of
11 magnitude more than the Internet router. Furthermore, more IP connections are made on a
12 daily basis than traditional telecom connections for phone calls in a month. As such, the
13 base technology necessary to replace an archaic telecom network is clearly available today.
14 The problem solved by the present invention is the method and system to implement a
15 network, which delivers telecom type services on a network built with new Internet
16 communication technology. This network not only replaces the current solution, but also
17 offers a more feature rich and cost effective alternative. Furthermore, the present invention
18 contains the ability for expansion via a scalable and extensible architecture to keep up with
19 the growth of the telecom industry. Thus, the present invention provides a replacement for
20 the technology of an outdated operating multibillion dollar infrastructure, which is cheaper,
21 more efficient and more reliable.

22 Furthermore, it is an overall object of the present invention to provide a VoIP
23 communication network that has a real time voice and data transmission profile and is
24 particularly useful in telephone communications implemented in a VoIP environment, such
as in a communication AppCenter, local exchange or other private telephone network. More

1 specifically, the present invention relates to a system and method relating to a private IP
2 communication network architecture facilitating audio, data, video, electrical, and logical
3 connections between two users.

4 One advantage of the present invention is to provide a protocol and network
5 architecture that allows communication between new and legacy components via a
6 translation interface module.

7 Another advantage of the present invention is to provide a user with a variety of
8 message types that can be optimized according to the type of data being sent, such as voice,
9 video, or data. The distributed network architecture also allows for the creation of a large
10 fault tolerant system that does not incur the performance and operational costs and
11 complexities associated with building a large monolithic system. Much like the differences
12 between Mainframe and local area network (LAN) computer environments. In the
13 distributed network architecture, transmitted data is identified and optimized by pooling data
14 into categorized data packet types allowing decisions to be made on how this data should be
15 handled and exchanged between servers along the various network types, such as a real time
16 packetized network or PSTN network, available to the user. Yet another advantage of the
17 present invention is the natural command and control infrastructure supplied by the
18 packetized network based architecture, namely, enhanced connectivity and scalability for
19 attached communication devices in a carrier grade network. While a packetized network
20 does not guarantee that all the advantages of an IP network will be available by default,
21 ATM is an example. The present invention utilizes the packetized network and IP protocol
22 to provide the ability address other devices without knowing where the device is or how to
23 get to it. Thereby separating network from application. The advantage of the system we are
24 describing is that this feature of IP networks has been sustained throughout the architecture.

1 In summary, the foregoing and other objects, advantages and features are achieved
2 with an improved communication network for use in connecting multimedia devices, such as
3 video, voice, data, other telephony devices, and the like to remote access points and
4 associated communication devices attached to those points, such as modems, telephones,
5 video displays, and network interface cards (NICs). Embodiments of the present invention
6 are particularly suitable for use with such telephony devices that are used in a typical local
7 exchange having one or more jacks or sockets designed to accommodate communication
8 devices. For example, a telephone having an RJ type connector that is inserted into the
9 socket or jack in such a way that the telephone is in communication with the network and
10 may selectively dial a second telephony device via the network.

11 In a preferred embodiment, the communication network maintains three layers for
12 each connection, more specifically a physical layer, a network layer, and a service layer.
13 The physical layer is created via the existing IP network, such as the Internet, a private IP
14 network, real time IP network, or combination thereof. A network layer generates a
15 connection that is logically assigned via participating network devices, and finally a service
16 layer generates logical connections necessary to run the desired application. By creating
17 three independent layers the communication network is given the advantages of a distributed
18 network along with the advantages of component specialization and optimization associated
19 with the physical, logical, and service layers. In designing the components of this new
20 communication network, the present invention employs, but is not limited to, elements from
21 Q.931 and SS7 as a basis for call control. A structured, scalable architecture is provided
22 through the defined components and their responsibility.

23 The network layer comprises at least one central arbitration server (CAS) or
24 gatekeeper, at least one communications engine (CE) or gateway, where both devices are

1 running the Internet Media Control Protocol (IMCP). The CAS is a fault tolerant set of
2 servers that act as gatekeepers on the network. The CAS is responsible for arbitrating
3 resource allocation, passing call control information and collecting billing records. The CE
4 is the network implementation of a voice-over IP (VoIP) gateway. This gateway is built to
5 use the IMCP protocol to take part in the private communication network. The network CE
6 can handle both customer based signaling, such as ISDN-PRI, and carrier signaling, such as
7 SS7. One example of an additional network layer device is a NetLink-IP device (C4). The
8 C4 provides network users with multiple phone lines and a persistent or continuous Internet
9 connection over a single data connection.

10 The service layer includes V-Link platform and services and the GateLink API layer.
11 The service layer brings intelligence to the network. V-Link is an example of a VoIP based
12 enhanced service. The network GateLink API offers access to the communication network
13 to third party software developers. The API gives a software abstraction to all the resources
14 on the network. Thereby enabling the creation of application modules, such as a voice
15 portal, a unified message service, and automatic call distribution (ACD) service.

16 In considering these factors, the following principles concerning new telephone
17 networks are observed by the present invention. First, support for legacy and current
18 communication standards and applications are available within the network. Second, the
19 network satisfies legacy, scalability, and reliability requirements of modern global business
20 models. Third, the network provides support for future communication features through
21 generalized buses and standardized communication protocols. Fourth, the network uses an
22 open architecture for third party vendor products and service extensions.

23 Merely rebuilding the telecom network in order to emulate the old legacy technology
24 is not in itself compelling, as the overall cost to replace the technology would be extreme.

1 Despite the fact that the new digital components for the new network architecture would be
2 less expensive and the overall network less cumbersome, the sheer size of the global
3 communications network makes an immediate universal replacement virtually impossible.
4 Instead the expectation is to replace the network in gradual steps over time. However,
5 various features and services can be effectively added to the new private IP communication
6 network in conjunction with many of the legacy components. Protocols, applications, and
7 devices within the network are designed without limiting them to current uses, but offering
8 an open adjustable, programmable network. The IMCP is extensible, allowing for additional
9 fields within a message or new messages. For example, additional fields include optional
10 data fields that a recipient need not understand (like SS7 information concerning a call to an
11 analog line) or even new private message blocks sent between devices or applications.

12 On the service layer, GateLink is the API an enhanced service is built on. The
13 GateLink API allows third party access to various network components. The API allows
14 software control of various functions, such as making a call, detecting DTMF tones, sending
15 a fax, recognizing speech, conferencing callers, and other functions. In this way the
16 developer need not bother with the underlying hardware just to implement an application.
17 The API approach makes the service layer of the present network architecture open and
18 expandable to support future services.

19 As is expected, a new communications network cannot initially exist by itself.
20 Before the entire telecom network is redeveloped, a transition period required, during
21 which the replacement network is capable of interconnecting with the older technology
22 networks. This allows for a seamless integration in support of current network services,
23 while allowing for a clear migration path for the industry. The Internet Media Control
24 Protocol (IMCP), which is the protocol at the base of the present invention, implements a

1 call control protocol based in part on the existing Q.931 standard, the same standard
2 implemented in other networks. Also, there is native support in the gateways (CE) for
3 protocols, such as ISDN-PRI and SS7. This means that without any additional modification
4 or customization, the network system of the present invention can connect into ISDN and
5 SS7 networks that presently exist today. The IMCP supports all current types of traffic
6 carried on a PSTN voice call: voice, fax, and modem. From an outside user standpoint, the
7 present invention looks like a traditional telecommunications network and yet due to the
8 inherent limitations of the PSTN network many advanced features of this new network are
9 not apparent. The inner workings that allow a lower cost and a higher functionality
10 communications network are transparent to the external user.

11 Despite the available connectivity, the method and system disclosed in the present
12 invention do not represent a communication network architecture intended to be grafted into
13 a larger switched network, but rather, constitutes an independent, stand alone, VoIP network
14 that features components with the scalability and versatility to surpass the service available
15 via existing traditional networks. Users have the ability to crossover to the existing
16 traditional telecommunication networks via translation modules, but this is a transitional
17 element of the network. The challenges presented by the legacy communication networks
18 during the transitional phase to a substitute or replacement communication network are
19 resolved through integration or incorporation with other VoIP networks via translation
20 modules. Moreover, the reliability of the present communication network is greatly
21 increased by taking advantage of dispersed, replicated and scalable components and
22 eliminating potential bottlenecks generally associated with PSTN, thereby lowering the
23 overall cost and complexity of the connecting network. Often the real time IP network is

24

1 able to facilitate these improvements through the mere addition of a server, as opposed to
2 the building of the huge monolithic devices required by the PSTN.

3 Additional objects and advantages of the invention will be set forth in the description
4 which follows, and in part will be obvious from the description, or may be learned by the
5 practice of the invention. The objects and advantages of the invention may be realized and
6 obtained by means of the instruments and combinations particularly pointed out in the
7 appended claims. These and other objects and features of the present invention will become
8 more fully apparent from the following description and appended claims, or may be learned
9 by the practice of the invention as set forth hereinafter.

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

BRIEF DESCRIPTION OF THE DRAWINGS

In order that the manner in which the above recited and other advantages and objects of the invention are obtained, a more particular description of the invention briefly described above will be rendered by reference to specific embodiments thereof that are illustrated in the appended drawings. Understanding that these drawing depict only typical embodiments of the invention and are not therefore to be considered to be limiting of its scope, the invention will be described and explained with additional specificity and detail through the use of the accompanying drawings in which:

Figure 1 illustrates an exemplary system that provides a suitable operating environment for the present invention;

Figures 2a, 2b, 2c, 2d, 2e, and 2f are a block diagrams of the control and logical connections illustrating various call states available via the packetized real time communication network architecture illustrated in Figure 1; and

Figure 3 is a timing diagram demonstrating a call.

1 **DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS**

2 Reference is first made to Figure 1, which graphically illustrates an exemplary
3 system or environment in which the present invention may be utilized or implemented.
4 Figure 1 is intended to be illustrative of potential systems that may utilize the present
5 invention and is not to be construed as limiting. Telephone calling devices 10a, 10b are
6 connected to central office (CO) 30a and 30b, respectively, via PSTN lines. CO 30a and
7 30b are electrically connected to the IP telephone gateways or communication engine (CE)
8 50a and 50b via PSTN connections or dedicated communication lines. Other telephony
9 devices 20a are connected to the network via CO 30c that is connected to CE gateway 50c.
10 CE gateway 50a, 50b, and 50c are all part of communication AppCenter 100, which may
11 include, among other things, telephony gatekeeper or central arbitration server (CAS) 40,
12 call detail record (CDR) database 45, network monitor 47, conference server 70, enhanced
13 service platform (V-Link) 60, GateLink server 82 for running communication applications
14 84, translation module 80, and communication proxy switch (C4P) 90. However, not all the
15 components are needed to be permanently associated with a single AppCenter. For
16 example, a single CAS can operational sustain many AppCenters. Furthermore, the CDR
17 database can be a central database that maintains all network activity. Even the conference
18 servers and V-Link services can be centrally located and provided to an AppCenter on
19 demand from a different location. Also the C4P and CE are not needed if the AppCenter
20 does not connect to C4s or PSTN customers respectively.

21 In the illustrated configuration, each module is connected to CAS 40 and is logically
22 interconnected during a call session with each other via a real time IP (RTIP) network 120.
23 The RTIP network 120 comprising: a private IP network, Internet, wireless IP network or
24 some combination of IP networks that enable suitable bandwidth for IP communication and

more preferably real time IP communication. All network traffic, including voice and real time applications, is preferably connected via RTIP network 120. C4P 90 combines data received from network connections generally through local digital switch (C4) modules 94a, 94b, 94c, and 94d. C4 modules 94 are similar to local digital telephone exchange centers. Each C4 may have multiple telephone connections 97a and 97b, multiple computer connections 97c, or other attached telephony devices 97d, such as a PBX. As previously discussed, the private IP network can be separated into three layers and their components: the network layer contains the IMCP protocol, CAS 40, CE 50, and other access devices, including the C4 device; the application layer contains the GateLink API , the AppLink platform and other related resources; and lastly, the service layer contains application modules for enhanced services, such as voice portals, unified messaging, ACD services, and other advanced communication applications. The advantages and functions of each of these layers and component modules are described in more detail below.

NETWORK LAYER

The network layer encompasses both a protocol and a hardware network. The physical network hardware comprises the routers, DS-1/ DS-3 links, PSTN connections, etc. The Internet Media Control Protocol (IMCP) contains a set of programming objects that control the resources available via the hardware network. For example, the protocol provides the mechanism for the network devices to communicate with each other, to reserve and control resources, and collect call detail records (CDR). The main network devices participating on the network layer are the central arbitration server (CAS) 40 and the communication engine (CE) 50. Other IMCP capable devices on the network include the C4 Proxy server (C4P) 90, GateLink 82, Conference servers 70, translation module 80, and V-Link servers 60.

1 In H.323 terms, CE 50 is, among other things, a digital gateway, and CAS 40 is a
2 gatekeeper. The CE 50 is the proxy between the existing telephony networks and an IMCP
3 interface with RTIP network 120. For example, the CE 50 may proxy as an ISDN-PRI
4 interface for the PSTN lines attached to CO 30. To connect to an SS7 network, one
5 network configuration uses a SS7 proxy (SS7P) to provide a SS7 signaling interface, while a
6 separate CE 50b provides payload conversion, such as encoding, compression, and other
7 IMCP formatting. CAS 40 is the connection control logic that maintains all network
8 connections, resource allocation and provides necessary billing information in the CDR. The
9 distributed network layer is intentionally designed with many redundancy components to
10 handle CAS 40 or IP connections during fail-over situations. However during normal
11 operations, a single CAS 40 easily carries millions of call setup requests per hour, due in
12 part to the distributed nature of the network and the low complexity of the IMCP protocol.
13 CAS 40 maintains resource allocation table, monitors network utilization, and tracks billing
14 information for the CDR, but has no specific knowledge or responsibility of the applications
15 using the resources. All other network devices support the IMCP protocol in the same
16 manner as CE 50. CAS 40 does not differentiate among the network devices (conferencing,
17 store & forward, client, text to speech, voice recognition, etc.). However, the IMCP
18 protocol can carry special messages required to separately control the function of the
19 different network devices.

20 **IMCP**

21 The Internet Media Control Protocol (IMCP) is at the lowest layer of the
22 architecture. This is the protocol that all devices and applications use to connect and
23 commune one with another. The IMCP protocol is designed to support the scalable and
24 feature rich VoIP network. The guiding design principles for the IMCP protocol include:

1 (1) reuse PSTN call setup protocol Q.931; (2) take advantage of the cost, performance,
2 pervasiveness, and scalability benefits supplied through the IP protocol; (3) support a
3 distributed, scalable architecture; (4) supply an open interface for other telecommunication
4 networks, such as the SS7 and H.323; and (5) support a feature rich network application and
5 device structure.

6 The IMCP protocol has two primary data activities: real time and control. The real
7 time portion or IMCP-data transfer is designed to carry the payload or the media packets
8 between two IMCP devices after a successful call setup. This would be, for example, voice,
9 fax, modem, silence, background noise, video, or other data types in the future. The control
10 portion of IMCP-call setup is designed to carry network events (DTMF and other tones),
11 applications, data, or private data and is illustrated in Figure 2a. IMCP-Call setup also
12 defines the messages required to setup a call between two IMCP endpoints. The relationship
13 of the real time portion to the control portion is illustrated in Figure 2b.

14 In addition to supporting the most important features required by high quality
15 telephony call setup, IMCP outperforms the basic requirements in some important aspects.
16 For example, IMCP features fast setup time where multiple events are handled in the same
17 message. Fast setup provides call setup times equivalent to or better than PSTN setup time.
18 The faster call setup is due in part to the fact that VoIP network signaling is only performed
19 at the end points and not at every switch along the call path. Additionally, the complexity
20 required from an IMCP terminal is minimal. Low complexity minimizes the load
21 requirements to process call setup, thereby allowing IMCP devices to be simple embedded
22 devices. In this respect IMCP is similar to the SIP protocol, in that IMCP messages are text
23 based and do not require special compilers or field allocation as in H.323 with ASN.1.
24 While the text-based approach does require higher bandwidth, the complexity reduction

1 during call setup outweighs this tradeoff. In fact, the overall higher bandwidth generates an
2 insignificant amount of data when compared to the real-time payload data being transmitted
3 in the overall scheme of the architecture. Another advantage of IMCP is the broad range of
4 PSTN support available. As the basic IMCP call setup procedure is compliant with the
5 Q.931 state machines, the interface to traditional PSTN networks is relatively
6 straightforward. This also improves the integration time of new servers using off the shelf
7 hardware and software components into the network. Yet another advantage of the text
8 based IMCP is the inherent support for new message types. As the IMCP message format is
9 text based, there are no coding limits and compatibility issues when new messages or
10 message types are added, additional text fields are easily ignored. Finally, IMCP is the base
11 protocol, so there are no lower layer protocols required. IMCP does not specify a lower
12 layer protocol. Unlike SIP, which runs on top of the HTTP protocol, or the H.323 that
13 requires an ASN.1 compiler and SSL, IMCP is simply integrated on top of the well-known
14 TCP/IP and UDP/IP protocols. This feature allows IMCP quick and efficient integration to
15 any device using a standard C compiler. Note that the IMCP call setup can be generalized to
16 any IMCP device, whether it is a PSTN gateway (CE) or a store and forward resource. It can
17 also be generalized to carry any type of media.

18 Another important feature supported by the IMCP protocol is the ability to transfer
19 connections among IMCP devices. This is done by a LinkLine message, which transfers the
20 real time connection to another IMCP device while keeping the control channel. This
21 feature is important when a network is required to support enhanced features beyond a
22 simple point-to-point connection. IMCP supplies the ability to transfer a call from one
23 platform to another while maintaining a control path to the originating platform. For
24 example, a calling card server will accept the first connection from the calling subscriber,

ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

1 will interact with the subscriber using common IVR techniques to authenticate the user and
2 to collect the destination number. The server will than initiate another call setup to the
3 destination number and once the call is accepted, will initiate a LinkLine request that will
4 transfer the real time connection between the subscriber and the destination number.
5 However, the server will maintain a control link to the originating CE 50 in order to play
6 “out of credit” warnings or to accept special requests from the subscriber using her or his
7 DTMF keypad. For the purpose of a connection control, the IMCP supports the
8 transmission of the DTMF detected signals over the control channel throughout the call
9 duration. This means that to support these features, the IMCP requires the originating CE 50
10 to detect and transfer the DTMF tones over the control channel.

11 A third feature supported by IMCP is the simple ability to add special messages
12 among IMCP components. CAS 40 will typically just transfer these messages but may also
13 decode them only for the purpose of special billing requests. Examples for these special
14 messages are conferencing, IVR, text to speech and voice recognition control. This requires
15 the IMCP components originating these messages to know the type of IMCP component
16 connected to by CAS 40.

17 The real time network 120 is built to support Real-time Transport Protocol (RTP)
18 and Internet Media Control Protocol – Real-time Transport (IMCP-RT). RTP itself does not
19 guarantee real-time delivery of data, but it does provide mechanisms for the sending and
20 receiving applications to support streaming data. Typically, RTP runs on top of the UDP
21 protocol, although the specification is general enough to support other transport protocols.
22 RTP has received wide industry support. As currently defined, RTP does not define any
23 mechanisms for recovering for packet loss. Such mechanisms are likely to be highly
24 dependent on the packet content and may be associated with the service layer and service

1 layer of the present invention. For example, for audio, it has been suggested to add low-bit-
2 rate redundancy, offset in time. For other applications, retransmission of lost packets may be
3 appropriate. (The H.261 RTP payload definition offers such a mechanism.) This requires no
4 additions to RTP. RTP probably has the necessary header information (like sequence
5 numbers) for some forms of error recovery by retransmission. In the present invention RTP
6 is supported equally to IMCP-RT.

7 The IMCP-RT is a lower overhead protocol designed to also carry information about
8 the data it carries. If more than one frame is destined for the same destination, the IMCP
9 layer will combine all the frames into a single UDP packet (or multiple packets in the case
10 of a large number of connection packets.) This can reduce the network bandwidth up to
11 40% and more in a real world environment. The frames within the packet are also labeled
12 with their content data type, such as voice, DTMF tones, facsimile, background noise,
13 digital data, modem, silence, or other data type. This labeling allows the end device to
14 process the frames without further analysis. A conferencing server would ignore packets
15 labeled as silence or background noise since there would be no need to add this data to a
16 conference call.

17 The control portion of the IMCP is a text-based protocol. All the data is sent as a
18 value-name pair. This allows for extensible messages that need not carry all the optional
19 fields if they are not used. It also allows for devices using different versions of the protocol
20 to use the same packets if the higher version device has backward compatibility. Higher-
21 level protocols, such as the call control, are implemented as a set of IMCP messages.

22 **CAS**

23 Every device in RTIP network 120 needs to be connected to CAS 40 via the IMCP in
24 order to participate in the RTIP network 120. Only one connection is made during the entire

1 uptime of the device. All calls, sessions, and ports for this devices are handled through the
2 same connection. The devices communicate with CAS 40 for log on, resource allocation,
3 and for data delivery. Data delivery may include delivering billing information, messages
4 for call control observed by CAS 40, and private messages transmitted for any purpose.
5 Data delivery may also involve private messages sent between devices and are passed
6 unobserved by the users. This strategy allows CAS 40 to handle network problems in a
7 reliable and efficient manner. For example, if a device, such as gateway, goes off the
8 network for any reason, only one device needs to be reconnected (as opposed to
9 reconnecting all the devices that are connected to it) since the remaining devices are still all
10 logged into CAS 40.

11 Reliability is fundamental to any carrier grade solution. As such, every component
12 within the present invention is redundant. The redundancy is loosely coupled without high
13 complexity and cost mirroring. Furthermore, high profile software based components, such
14 as CAS 40, have seamless redundancy. For example, if the network cable from CAS 40 is
15 unplugged the system will not lose a single billing record nor will the disconnection affect a
16 call in progress. This occurs in part, because all the network devices will switch to a
17 redundant CAS. Once the cable is plugged back in, CAS 40 uploads the billing records and
18 continues operating. The calls in progress are not affected because payload is transmitted
19 directly between the end units that carry the voice or other data packets between the
20 connected users participating in the call. Only the ICMP control messages are routed via
21 CAS 40, so that calls in progress can continue uninterrupted by a disconnected CAS. The
22 originating and receiving units switch to the redundant CAS and deliver what information
23 they have about their current state. The control data being sent to the redundant CAS is not
24 as time sensitive as the voice data and can absorb any delay introduced by a fail over. In one

1 configuration of an AppCenter, multiple CAS units are available so that the device control
2 lines can be transferred to an operating CAS in fail-over situations.

3 Thus the present invention is designed for reliability on multiple layers. Separating
4 functionality, like application from network, allows for a robust scalable architecture. For
5 example, on the service layer, AppLink is built to recover from the database failure. If the
6 database connection is lost, the AppLink server will reconnect without even returning the
7 application an error, thereby ensuring the caller an uninterrupted telephone session.

8 CAS 40 is similar to a gatekeeper in an H.323 network or a SCP in an SS7 network.
9 However, the primary responsibilities of CAS 40 are to keep track of resource utilization,
10 pass information between devices, collect billing information and collect and deliver
11 monitoring information about the devices it manages. As such CAS 40 is the perfect device
12 from which a network monitor 47 or billing database 45 can obtain their information.

13 CAS 40 is a distributed application that enables resource replication to enhance the
14 overall network system by adding new devices, applications, and components. Replication
15 allows the singular network device performance to be amplified by replicating a device on
16 the network. For example, using currently available hardware configurations a single CAS
17 can handle at least five million calls an hour, a single AppLink server can control at least
18 one thousand enhanced service sessions, where the typical delay from the time a network
19 event occurs until it is visible on a maintenance console or network monitor 47 is about fifty
20 milliseconds, and the typical time to reload a route table containing all routes for the entire
21 communication system while CAS 40 is running at a high load stress is about three seconds.

22 In an alternative configuration, CAS 40 is also a distributed application that enables
23 resource replication to enhance the overall network capacity. Also, replication allows a

1 distributed application on a network to get more than one thousand enhanced service
2 sessions simply by adding another GateLink server.

3 As a result of the network layer architecture and the IMCP, additional operational
4 features are available to the network. The two most important features are CDR collection
5 and system monitoring. Both of these features are directly related to the fact the all IMCP
6 messages are routed through the CAS. The CAS sees all the call control messages and can
7 populate all CDR by default with: originating and terminating number; CE line; and trunk
8 group and call start, answer, and end. In addition the CAS allows for extensible CDR
9 allowing the application to add any fields needed to completely describe a call like type of
10 service for instance. Also CAS allows an application to "group" CDR together with a "key"
11 to allow later bill creation to present a complex session like a conference call in a way the
12 customer will understand. System monitoring is possible since the CAS has all the states of
13 all the lines of the devices in the network. The CAS contains information if a line is active
14 and what other device it is connected to and for how long. Depending on the application and
15 system management tools, this architecture can be extended to provide carrier grade services
16 in both a scalable and reliable manner.

17 **CE**

18 Communication engine (CE) 50 is a VoIP gateway in the private IP network. The
19 CE uses IMCP to communicate. In one embodiment, CE 50 is an industrial PC with enough
20 network cards and DSP resources to handle ten T1 lines worth of telephone calls. Future
21 plans for an embedded version and larger, compact PCI version will enable CE 50 to carry
22 more calls and be more reliable. CE 50 acts like a gateway from an information poor PSTN
23 signal to an information rich IMCP network. CE 50 can not only compress voice data but
24 identifies and categorizes the data. Packets are labeled as voice, silence, background noise,

1 DTMF tones, fax, or modem. In addition, fax traffic is demodulated and the raw T.30
2 information is sent in the packets. This allows other devices and applications to manipulate
3 the data without the need of further DSP analysis. IVR (Integrated Voice Response)
4 systems can detect a DTMF tone by checking for DTMF packets. A voice recognition
5 server can detect when an end of a word occurs by the silence packets. Yet another
6 application provides a store and forward fax solution, which uses the T.30 information to
7 create an IP based fax service. More specifically, the CE recognizes the PCM or modulate
8 wave signal from the facsimile device and repackages the information into modules, such as
9 V.17, FSK, and CNG for transmission to a GateLink server running the IP based fax service
10 application. An application running a fax service via the GateLink API is then able to
11 access a T.30 state machine for operations, such as "send fax" and "receive fax", on the
12 GateLink server without needing to interpret PCM. The fax service application would then
13 be able to generate .TIFF, .JPG, .GIF, or other similar graphics file types of the original fax.
14 As described the CE must repackage the PCM data through demodulation into new modules
15 without performing any DSP operations. The DSP operations are accomplished on the
16 GateLink server via the fax service application. The CE, in short, is the electrical muscle for
17 the brains, which reside on the service layer.

18 **C4**

19 The NetLink-IP (C4) 95 is an example of the next generation of access devices
20 attached to RTIP network 120. C4 95 allows the network the ability to offer a customer
21 multiple phone lines and a persistent Internet connection over a single data line connection.
22 C4 95 is the first integrated CPE to connect to a VoIP network and deliver all the services
23 that previously required the use of class 5 switches. C4 95 delivers the intelligence and
24 benefits of the previously described RTIP network 120 all the way to the consumer.

1 In summary, the user is connected to RTIP network 120 from the time they pick up
2 their handset without having the traditional telecom network to limit the control and features
3 that RTIP network 120 can provide to the user. RTIP network 120 is designed as a complete
4 replacement for the traditional telecom networks. Thus, the new C4 architecture allows for
5 this network to connect to the traditional networks and allows for an upgrade path. The
6 design of this architecture is robust and scalable so that this network can introduce new
7 features and functionality while preserving the quality of traditional networks.

8

9

SERVICE LAYER

10 The service layer takes advantage of the network components in the RTIP network
11 120 to provide an environment for building a high performance, scalable and feature rich
12 communication network. As the underlying network to the API already handles many duties
13 of a telecom application, the service layer needs to worry only about the application itself.
14 CAS 40, for example, handles resource allocation and locking issues, the ICMP protocol and
15 GateLink API handle the complexities of manipulating resources in the network, and the
16 CEs handle pre-digestion of the signal, relieving the application of any need for a DSP
17 resource.

18

19

20

21

22

23

24

 An example of how the service layer interacts with the network layer can be seen
from the following description of a one number call. A one number call is the ability of a
caller to dial a single number and have the one number application reach the subscriber at
multiple numbers at once. The initiating caller will call a number assigned to the
subscriber's one number service. This call will come into CE 50. CAS 40 will, based on the
called number, route the call to an appropriate AppLink server and lock resources on both
CE 50 and the AppLink server. The GateLink API will handle all the ICMP call controls to

receive this call via "Wait for Call" and "Answer Call" API calls. With the call information delivered via the IMCP, the AppLink server will identify which user is being called and play the appropriate greeting. The "Play Prompt" API delivers saved frames from the caller via the IMCP-RT protocol. CE 50 will convert these saved frames into speech and the caller will hear the greeting. "Get DTMF digit" will wait for the caller to press a designated number to locate the subscriber. Separate "Make Call" API calls will call the subscriber. "Play Prompt" module will play a greeting of the caller previously recorded with "Record Prompt" and "get DTMF digit" module will await a response from the subscriber indicating that he is ready to receive the caller. The application will now have two sessions: one with CE 50a with the caller, and one with the subscriber. The applications will "Link Line" these two sessions, allowing the IMCP-RT packets to travel directly between the two CEs. The IMCP control session remains the same even though the RT packet paths have changed. All billing information specific to the application, such as what type of phone number did the subscriber answer, is passed to the CAS 40 and recorded. In addition, the call records for both the caller and subscriber contain a key indicating that they belong to the same session. It is noteworthy that the API "get DTMF digit" does not actually look at the signal or the Real Time packets. DTMF tones sent by the caller are identified by CE 50a and are also sent as messages via IMCP. The application can receive DTMF tones, even once the caller and subscriber are connected and IMCP-RT packets are transferred directly between CE 50s, thereby enabling DTMF direction across the lines.

Reference is next made to Figure 2, block diagrams of the method and system for interconnecting a private IP communication network. Figure 2a represents the control paths that are established between various network devices and central arbitration server (CAS) 40 as the network devices "log in" to the network. The continuous control line structure is

1 illustrated for gateways 50a and 50b, the conference server 70, the V-Link enhanced service
2 platform 60, and the CAS 40. These control lines determine whether or not a call may be
3 connected and contain information concerning the phone conversation such as billing
4 information without burdening the direct connection between the devices.

Figure 2b demonstrates a variety of potential real time data paths that may exist between network devices. For example, gateway 50a may be connected directly with gateway 50b, or indirectly connected via the V-Link enhanced service platform 60, or the conference server 70. The real time data paths illustrated in Figure 2b represent selective network connections and selective logical connections between the network devices, while the control path connections as depicted in Figure 2a are full time connections between the network devices.

Figure 2c illustrates the first step in creating a special service call using enhanced services V-Link server 60. Origination gateway 50a and destination gateway 50b connect to V-Link server 60 via data and logical control lines. V-Link server 60 is unique in its methodology and flexibility when interacting with other network devices. For example, a CE would deliver encoded packets from the PSTN connection, but V-Link server 60 delivers packets from a disk that play a greeting and instruct the origination and destination user. User input is received from the origination and destination gateways 50a and 50b via DTMF messages and user messages that are recorded to disk or memory, in essence building an IVR (Integrated Voice Response) environment. Based on DTMF input from the caller requesting to connect to the subscriber, the V-Link platform places an outgoing call in attempt to reach the subscriber to termination CE 50b (CAS decides this based on the telephone number of the subscriber). When the subscriber answers, the call is considered “Connected”. There are now two connections to V-Link server 60: the origination caller

1 connection and the subscriber connection and the connections between originator and
2 subscriber remain active until the end of the call in one form or another.

3 With reference to Figure 2d, once the destination subscriber answers the phone and
4 accepts the originating call, there is a need to connect the two data lines. In a normal call the
5 data path would follow the logical control path, that is, origination gateway 50a connects to
6 destination gateway 50b and the “voice” data is also sent from origination gateway 50a to
7 destination gateway 50b. But in the conference call situation, the network handles the call
8 differently. Namely, V-Link server 60 will use the “LinkLine” command via CAS 40 to tell
9 origination gateway 50a and destination gateway 50b to deliver “voice” data to each other
10 while still maintaining a control path to V-Link server 60. So in a logical sense both the
11 originating caller and destination subscriber are still connected to V-Link server 60, but their
12 voice data path is redirected to each other. This allows V-Link server 60 to maintain
13 supervision (both line and DTMF) of the call without having to route all the “voice” data
14 through V-Link server 60.

15 This comes in handy when the destination subscriber decides to create a simplified
16 conference call as illustrated in 2e and 2f. A digit sequence, for example “00” alerts V-Link
17 server 60 that the subscriber needs access to the system and uses “LinkLine” to connect both
18 data path calls back to V-Link server 60. The caller will receive packets for music on hold
19 and the subscriber will be in the IVR system associated with V-Link server 60. A menu
20 system within the IVR system instructs the subscriber concerning the available services,
21 including instructions on how to build a conference call. As a result of the subscribers input,
22 the system in Figure 2e creates two calls to conference server 70 via V-Link server 60. The
23 first call to conference server 70 creates a new conference session identifier and the second
24 call delivers the session identifier in a user field via IMCP, thereby placing both calls in the

1 same conference. These calls remain for the duration of the conference call. Then V-link
2 will use the “LinkLine” command to connect the data paths from V-Link to the Conference
3 server, as depicted in Figure 2f.

4 Reference is now made to Figures 1 and 3. Figure 3 illustrates a call flow chart
5 indicating the process of establishing a phone call between a PSTN telephone user 10a to a
6 second PSTN telephone user 10b. In this situation, a call is placed from the PSTN
7 origination point 10a, the call travels through the CO 30a and arrives at the communication
8 engine (CE) or gateway 50a.

9 Figure 3 describes one embodiment of the call flow during an IMCP call setup
10 session between two CEs (Gateways) 50 and CAS 40. When a first originating CE gateway
11 50a receives a call setup request from an attached PSTN line user 10a, the originating CE
12 gateway initiates a “LockLine” signal request with enough calling information to CAS 40 to
13 determine which terminating CE gateway 50 would be best suited to carry the call. Calling
14 information includes information such as the destination phone number and the requested
15 bandwidth. A LockLine signal request requires a network resource with specific
16 parameters, such as destination phone number. CAS 40, based in part on its dynamic
17 routing tables, determines the line availability in the closest available termination CE
18 gateway to the call destination. The CAS allocates and acknowledges the resource
19 availability with a “LockLineAck” signal message to the originating CE gateway, along with
20 information corresponding to the termination CE gateway. For example, CAS 40 can
21 transfer the IP address of the termination CE gateway to the originating CE gateway,
22 enabling the network to create the real time connections to carry the media information
23 directly between both IMCP endpoints. In turn, the originating CE gateway sends a
24 “Proceeding” signal to the PSTN originating device. The CAS also marks the ports on both

1 originating and termination CE gateways as locked, making them busy or inaccessible to
2 subsequent calls.

3 This resource acknowledgement triggers a call request or “MakeCall” signal that is
4 monitored by CAS 40 from the originating CE gateway 50a to the termination CE gateway
5 50b. Using this call request signal, the originating CE 50a can force or suggest the call
6 parameters for the call. The termination CE 50b then initiates call “Setup” signal to connect
7 with the PSTN destination. The PSTN destination acknowledges the “Setup” signal with a
8 “Proceeding” signal followed by an “Alerting” signal. The termination CE 50b forwards the
9 Alerting signal, monitored by CAS 40, along with additional call information to the
10 originating CE 50a. The originating CE 50a forwards the Alerting signal to the originating
11 PSTN subscriber. The while the timing diagram illustrated in Figure 3 illustrates an
12 accepted call, the call response signals may be one of a set of possible responses based on
13 the success or failure in making the call. For example, “AcceptCall” may produce an
14 Alerting signal while “ConnectCall” will indicate that the destination is connected or
15 “ClearCall” is used when the line is busy or unavailable.

16 Following the Alerting signal, a “Connect” signal is transmitted from the destination
17 PSTN to the termination CE gateway. The connecting signal is monitored by the CAS and
18 then forwarded from the termination CE directly to the originating CE, which then forwards
19 the Connect signal to the originating PSTN call point. Upon the end of a call, the “Clear”
20 signal is sent from the originating PSTN to the originating CE. The originating CE forwards
21 this “Clear” signal directly to the termination CE gateway, which then forwards the clear
22 signal to the destination PSTN device. The destination PSTN device then transmits a “Clear
23 Acknowledge” signal to the termination CE. The termination CE then transmits the clear
24 acknowledge signal to the originating CE, which forwards the “Clear Acknowledge” signal

1 to the originating PSTN. In all cases, the CAS monitors the call control signals so that CAS
2 can accurately allocate network resources. The “Clear” signal is illustrated as being
3 originated from the originating PSTN device and the “Clear Acknowledge” signal is
4 illustrated as being generated by the destination PSTN device. However, the “Clear” or
5 “Clear Acknowledge” signals may be originated from either the origination or destination
6 device, depending on who ends the call first.

7 Resource allocation is the responsibility of CAS 40. If there is more than one CE 50
8 that could handle the termination, CAS 40 decides where to send the call. CAS 40 uses a
9 database table that maps telephone number ranges from an NPA all the way down to a
10 specific phone number for various end devices. If the devices are the same priority, the call
11 load will be equally distributed, or if a different priority, the higher priority will be used
12 until they are full, allowing overflow, class of service, failure bypass, and least cost routing.
13 Least cost routing chooses the cheapest path for the data to be transmitted. Class of service
14 assigns a prioritization to certain customer data types. For example, a customer “paying”
15 for data payload would take priority over a “free access” data payload. Another routing
16 method used to improve connections between end devices is failure bypass routing,
17 commonly used to avoid portions of the network that are either not performing or are
18 performing poorly, such as overloaded network sections that function below a user specified
19 response performance parameter. Once the originating CE gateway receives the necessary
20 information about the termination CE gateway from the CAS, the CAS observes certain call
21 control messages that are passed between connected end devices via the IMCP. The CAS
22 maintains call state information for each port, whether the port is idle, alerting, or connected.
23 This call related port data is used for network monitoring and for billing information.

24

1 Billing information about the call is split into two parts: a base record and fields
2 associated with the base record. The base record includes all the basic network level CDR
3 information, such as call start-answer-end times, ports, machines, etc. Because this is a
4 distributed application with many devices working together to deliver a service, time stamps
5 on these call records are kept with millisecond accuracy. CAS 40 additionally implements a
6 powerful concept of "fields" into its CDR. This allows communication applications via the
7 IMCP to deliver to CAS 40 any number of additional fields on a per record basis. Enabling
8 CAS 40 to collect billing information for any application without having to anticipate the
9 application. A fax on demand application, for instance, could collect a list of documents
10 each user sent. A unified messaging or voicemail application could bill by the number of
11 messages the user listened to. These generic fields allow applications to use the high
12 performance and reliability of CAS 40 without sacrificing information concerning related
13 billing records. Since many applications, such as conferencing and multiple call legs
14 associated with a single session (like a one number call), require more than one call record
15 to record all legs of a call, CAS also has the ability to group call records using a key. All of
 the legs of a conference call could share the same key allowing a simpler bill to be sent.

 The present invention may be embodied in other specific forms without departing
 from its spirit or essential characteristics. The described embodiments are to be considered
 in all respects only as illustrative and not restrictive. The scope of the invention is,
 therefore, indicated by the appended claims rather than by the foregoing description. All
 changes that come within the meaning and range of equivalency of the claims are to be
 embraced within their scope.

23 What is claimed and desired to be secured by United States Letters Patent is:

24